



OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/939,408A

DATE: 01/23/2002
TIME: 17:03:26

Input Set : A:\500nscseq.txt
Output Set: N:\CRF3\01232002\I939408A.raw

Does Not Comp
Corrected Discrete

3 <110> APPLICANT: Yoshida, Roberta
4 Kootstra, Anna
7 <120> TITLE OF INVENTION: Phenylalanine Ammonia Lyase Polypeptide and
8 Polynucleotide Sequences and Methods of Obtaining and
9 Using Same
11 <130> FILE REFERENCE: 29479/500NSCA
13 <140> CURRENT APPLICATION NUMBER: US 09/939,408A
17 <141> CURRENT FILING DATE: 2001-08-24
19 <150> PRIOR APPLICATION NUMBER: US 09/624,693
22 <151> PRIOR FILING DATE: 2000-07-24
23 <150> PRIOR APPLICATION NUMBER: PCT/US01/23270
26 <151> PRIOR FILING DATE: 2001-07-24
29 <160> NUMBER OF SEQ ID NOS: 30
31 <170> SOFTWARE: PatentIn Ver. 2.0

ERRORED SEQUENCES

1706 <210> SEQ ID NO: 21
1707 <211> LENGTH: 726
1708 <212> TYPE: PRT
1709 <213> ORGANISM: Artificial Sequence
1711 <220> FEATURE:
1712 <221> NAME/KEY: SITE
1713 <222> LOCATION: (12)..(719)
1714 <223> OTHER INFORMATION: "Xaa"means any amino acid;"Xaa"means no consensus at that
position
1716 <223> OTHER INFORMATION: Description of Artificial Sequence: Consensus of
1717 SEQ ID NOS: 13, 17, and 19
1719 <400> SEQUENCE: 21
W--> 1721 Met Ala Pro Ser Leu Asp Ser Ile Ala Thr Ser Xaa Ala Asn Gly Xaa
1722 1 5 10 15
W--> 1724 Xaa Asn Gly Xaa His Ala Ala Xaa Xaa Ala Ser Xaa Xaa Xaa Xaa Xaa
1725 20 25 30
W--> 1727 Xaa Xaa Xaa Xaa Ala Xaa Ala Gly Ser Xaa Leu Pro Thr Thr Xaa Xaa
1728 35 40 45
W--> 1730 Thr Gln Leu Asp Ile Val Glu Xaa Xaa Leu Ala Asp Pro Xaa Thr Asp
1731 50 55 60
W--> 1733 Asp Xaa Xaa Glu Leu Asp Gly Tyr Ser Leu Thr Leu Gly Asp Val Val
1734 65 70 75 80
W--> 1736 Gly Ala Ala Arg Lys Gly Arg Xaa Val Arg Val Xaa Asp Ser Asp Glu
1737 85 90 95
W--> 1738

Ile Arg Xaa Lys

more over

E--> 1739

100

105

110

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/939,408A

DATE: 01/23/2002
 TIME: 17:03:26

Input Set : A:\500nscseq.txt
 Output Set: N:\CRF3\01232002\I939408A.raw

```

W--> 1741  Xaa Asn Ser Val Tyr Gly Val Thr Thr Gly Phe Gly Ser Ala Asp
          115 120 125
1742  Thr Arg Thr Glu Asp Ala Ile Ser Leu Gln Lys Ala Leu Leu Glu His
          130 135 140
1744  Gln Leu Cys Gly Val Leu Pro Thr Ser Xaa Asp Ser Phe Xaa Leu Gly
          145 150 155 160
W--> 1747  Arg Gly Leu Glu Asn Ser Leu Pro Leu Glu Val Val Arg Gly Ala Met
          165 170 175
1748  Thr Ile Arg Val Asn Ser Leu Thr Arg Gly His Ser Ala Val Arg Leu
          180 185 190
1750  Val Val Leu Glu Ala Leu Thr Asn Phe Leu Asn His Gly Ile Thr Pro
          195 200 205
1751  Ile Val Pro Leu Arg Gly Thr Ile Ser Ala Ser Gly Asp Leu Ser Pro
          210 215 220
1753  Leu Ser Tyr Ile Ala Ala Ala Ile Thr Gly His Pro Asp Ser Lys Val
          225 230 235 240
1754  His Val Xaa His Glu Gly Xaa Glu Lys Ile Met Xaa Ala Arg Glu Ala
          245 250 255
W--> 1765  Ile Ala Leu Phe Gly Leu Glu Pro Val Val Leu Gly Pro Lys Glu Gly
          260 265 270
1766  Leu Gly Leu Val Asn Gly Thr Ala Val Ser Ala Ser Met Ala Thr Leu
          275 280 285
1768  Ala Leu His Asp Ala His Met Leu Ser Leu Leu Ser Gln Ala Leu Thr
          290 295 300
1769  Ala Leu Thr Val Glu Ala Met Val Gly His Ala Gly Ser Phe His Pro
          305 310 315 320
1771  Phe Leu His Asp Val Thr Arg Pro His Pro Thr Gln Ile Glu Val Ala
          325 330 335
1772  Arg Asn Ile Arg Thr Leu Leu Glu Gly Ser Xaa Phe Ala Val His His
          340 345 350
W--> 1783  Glu Glu Glu Val Lys Val Lys Asp Asp Glu Gly Ile Leu Arg Gln Asp
          355 360 365
1784  Arg Tyr Pro Leu Arg Thr Ser Pro Gln Trp Leu Gly Pro Leu Val Ser
          370 375 380
1786  Asp Met Ile His Ala His Ala Val Leu Ser Leu Glu Ala Gly Gln Ser
          385 390 395 400
1787  Thr Thr Asp Asn Pro Leu Ile Asp Val Glu Asn Lys Xaa Thr His His
          405 410 415
W--> 1795  Gly Gly Asn Phe Gln Ala Ser Ala Val Xaa Asn Thr Met Glu Lys Thr
          420 425 430
1796  Arg Leu Ala Leu Ala Leu Ile Gly Lys Leu Asn Phe Thr Gln Leu Thr
          435 440 445
1801  Glu Met Leu Asn Ala Gly Met Asn Arg Gly Leu Pro Ser Cys Leu Ala
          450 455 460
1802  Ala Glu Asp Pro Ser Leu Ser Tyr His Cys Lys Gly Leu Asp Ile Ala
          465 470 475 480
1804  Ala Ala Ala Tyr Thr Ser Glu Leu Gly His Leu Ala Asn Pro Val Thr
          485 490 495
1805  Thr His Val Gln Pro Ala Glu Met Gly Asn Gln Ala Val Asn Ser Leu
          495
1807  1810 1811 1813

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/939,408A

DATE: 01/23/2002
TIME: 17:03:26

Input Set : A:\500nscseq.txt
Output Set: N:\CRF3\01232002\I939408A.raw

```

1814           500           505           510
1816 Ala Leu Ile Ser Ala Arg Arg Thr Ala Glu Ala Asn Asp Val Leu Ser
1817           515           520           525
1819 Leu Leu Leu Ala Thr His Leu Tyr Cys Val Leu Gln Ala Val Asp Leu
1820           530           535           540           560
W--> 1822 Arg Ala Met Glu Phe Glu Phe Lys Lys Gln Phe Xaa Pro Xaa Xaa Xaa
1823           545           550           555           570           575
W--> 1825 Xaa Leu Leu Xaa Gln His Phe Gly Xaa Xaa Xaa Thr Xaa Xaa Xaa Xaa
1826           565           580           585           590           595
W--> 1828 Xaa Xaa Glu Leu Xaa Xaa Lys Val Xaa Lys Xaa Leu Xaa Lys Arg Leu
1829 Glu Gln Thr Asn Ser Tyr Asp Leu Glu Pro Arg Trp His Asp Ala Phe
1831           595           600           605
1832 Ser Xaa Ala Thr Gly Thr Val Val Glu Xaa Leu Ser Ser Xaa Xaa Xaa
W--> 1834           610           615           620           635           640
1835 Xaa Xaa Val Ser Leu Ala Ala Val Asn Ala Trp Lys Val Ala Xaa Ala
W--> 1837           625           630           635           640           655
1838 Glu Lys Ala Ile Ser Leu Thr Arg Xaa Val Arg Xaa Xaa Phe Trp Xaa
W--> 1840           645           650           655
1841 Ala Pro Ser Ser Ser Ser Pro Ala Leu Xaa Tyr Leu Ser Pro Arg Thr
W--> 1843           660           665           670           685
1844 Arg Val Leu Tyr Ser Phe Val Arg Glu Glu Leu Gly Val Lys Ala Arg
1846           675           680           685
1847 Arg Gly Asp Val Phe Leu Gly Lys Gln Glu Val Thr Ile Gly Ser Asn
1849           690           695           700           715           720
1850 Val Ser Arg Ile Tyr Glu Ala Ile Lys Ser Gly Arg Ile Asn Xaa Val
W--> 1852           705           710           715           720
1853 Leu Val Lys Met Leu Ala
1855           725
1856

```

```

1919 <210> SEQ ID NO: 28
1920 <211> LENGTH: 2741
1921 <212> TYPE: DNA
1922 <213> ORGANISM: Rhodotorula graminis
W--> 1923 <220> FEATURE:
1924 <221> NAME/KEY: modified_base
1925 <222> LOCATION: (6)
1926 <223> OTHER INFORMATION: Description of modified_base: m = a or c
W--> 1927 <220> FEATURE:
1928 <221> NAME/KEY: exon
1929 <222> LOCATION: (2008)..(2586)
1930 <223> OTHER INFORMATION:
W--> 1931 <220> FEATURE:
1932 <221> NAME/KEY: exon
1933 <222> LOCATION: (1822)..(1947)
1934 <223> OTHER INFORMATION:
W--> 1935 <220> FEATURE:
1936 <221> NAME/KEY: exon
1937 <222> LOCATION: (1587)..(1748)
1938 <223> OTHER INFORMATION:

```

DATE: 01/23/2002
TIME: 17:03:26

Input Set : A:\500nscseq.txt
Output Set: N:\CRF3\01232002\I939408A.raw

W--> 1939 <220> FEATURE:
1940 <221> NAME/KEY: exon
1941 <222> LOCATION: (1365)..(1529)
1942 <223> OTHER INFORMATION:

W--> 1943 <220> FEATURE:
1944 <221> NAME/KEY: exon
1945 <222> LOCATION: (961)..(1295)
1946 <223> OTHER INFORMATION:

W--> 1947 <220> FEATURE:
1948 <221> NAME/KEY: exon
1949 <222> LOCATION: (449)..(880)
1950 <223> OTHER INFORMATION:

```

W--> 1951 <220> FEATURE:
1952 <221> NAME/KEY: exon
1953 <222> LOCATION: (1)..(361)
1954 <223> OTHER INFORMATION:

```

W--> 1955 <220> FEATURE:
1956 <221> NAME/KEY: terminator
1957 <222> LOCATION: (2587)..(2589)
1958 <223> OTHER INFORMATION:

W--> 1959 <220> FEATURE:
1960 <221> NAME/KEY: Intron
1961 <222> LOCATION: (1948)..(2007)
1962 <223> OTHER INFORMATION:

W--> 1963 <220> FEATURE:
1964 <221> NAME/KEY: Intron
1965 <222> LOCATION: (1749)..(1821)
1966 <223> OTHER INFORMATION:

W--> 1967 <220> FEATURE:
1968 <221> NAME/KEY: Intron
1969 <222> LOCATION: (1530)..(1586)
1970 <223> OTHER INFORMATION:

W--> 1971 <220> FEATURE:
1972 <221> NAME/KEY: Intron
1973 <222> LOCATION: (1296)..(1364)
1974 <223> OTHER INFORMATION:

W--> 1975 <220> FEATURE:
1976 <221> NAME/KEY: Intron
1977 <222> LOCATION: (881)..(960)
1978 <223> OTHER INFORMATION:

W--> 1979 <220> FEATURE:
1980 <221> NAME/KEY: Intron
1981 <222> LOCATION: (362)..(448)
1982 <223> OTHER INFORMATION:

1982 <223> OTHER INFORMATION:
W--> 1983 <400> SEQUENCE: 28
1984 atg gcm cct tcc ttg gac tgc ctc gcc acc acg ctc gcc aac ggc ttt
asp ser leu ala thr thr leu ala asn gly phe 48

1984 atg gcm ccc tcc ccg ggc
1985 Met Ala Pro Ser Leu Asp Ser Leu Ala Thr Thr Leu Ala
F--> 1986 10 Ser Ala Ala Gly Pro Thr
15 insert
hard return ← acc aac ggc tcg c

E--> 1986 10
1 5
W--> 1987 Thr Asn Gly Ser His Ala Ala Pro Thr Lys Ser Ala Ala Gly Pro Thr

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/939,408A

DATE: 01/23/2002
TIME: 17:03:26

Input Set : A:\500nscseq.txt
Output Set: N:\CRF3\01232002\I939408A.raw

W--> 1988	20	25	30	144
1989	tcg gct ctc cgc cgc acg ccc ggc ctc gat ggc cac gcc gcg cac cag			
1990	Ser Ala Leu Arg Arg Thr Pro Gly Leu Asp Gly His Ala Ala His Gln			
W--> 1991	35	40	45	192
1992	tcg cag ctc gag atc gtg cag gag ctc ctc agc gac ccc acc gac gac			
1993	Ser Gln Leu Glu Ile Val Gln Glu Leu Leu Ser Asp Pro Thr Asp Asp			
W--> 1994	50	55	60	240
1995	gtc gtc gag ctc agc ggg tac agc ctc acc gtc cgt gac gtt gtc ggc			
1996	Val Val Glu Leu Ser Gly Tyr Ser Leu Thr Val Arg Asp Val Val Gly			
W--> 1997	65	70	75	288
1998	gcc gcc cgc aag ggg cgc agg gtc cgc gtc cag aac gac gac gag atc			
1999	Ala Ala Arg Lys Gly Arg Arg Val Arg Val Gln Asn Asp Asp Glu Ile			
W--> 2000	85	90	95	336
2001	cgc gca cgc gtc gac aag agc gtc gac ttc ctc aag gcc cag ctt cag			
2002	Arg Ala Arg Val Asp Lys Ser Val Asp Phe Leu Lys Ala Gln Leu Gln			
W--> 2003	100	105	110	381
2004	aac tcg gtc tac gga gtc acc acg g tgcgttcga gacgagaggc			
2005	Asn Ser Val Tyr Gly Val Thr Thr			
W--> 2006	115	120		441
2007	ggaaatctcg ggatgccgca gcgctgaacg ctgacactcg cttggacggc tgccgcggtc			489
2008	ttgcagg gt ttc ggt ggc tcg gcc gac acg agg act gag gat gca gtc			
2009	Gly Phe Gly Gly Ser Ala Asp Thr Arg Thr Glu Asp Ala Val			
W--> 2010	125	130		537
2011	agc ctc cag aag gcg ctc atc gag cac cag ctc tgc ggc gtg acg ccg			
2012	Ser Leu Gln Lys Ala Leu Ile Glu His Gln Leu Cys Gly Val Thr Pro			
W--> 2013	135	140	145	585
2014	acg tcc gtc tcg tcc ttc agc gtc gga cgc ggc ctc gag aac acg ctt			
2015	Thr Ser Val Ser Ser Phe Ser Val Gly Arg Gly Leu Glu Asn Thr Leu			
W--> 2016	155	160	165	633
2017	ccg ctc gag gtc gtc cgc ggc gcc atg gtc atc cgc gtc aac tcg ctc			
2018	Pro Leu Glu Val Val Arg Gly Ala Met Val Ile Arg Val Asn Ser Leu			
W--> 2019	170	175	180	681
2020	acg cgt ggc cac tcg gcc gtc cgc ctc gtc gtc ctt gag gcg ctc acc			
2021	Thr Arg Gly His Ser Ala Val Arg Leu Val Val Leu Glu Ala Leu Thr			
W--> 2022	185	190	195	729
2023	aac ttc ttg aac cac cgc atc acg ccc atc gtc ccc ctc cgc ggc tcc			
2024	Asn Phe Leu Asn His Arg Ile Thr Pro Ile Val Pro Leu Arg Gly Ser			
W--> 2025	200	205	210	777
2026	atc tcg gcg tcg ggc gac ctc agc ccg ctc tcg tac atc gcc ggc gcc			
2027	Ile Ser Ala Ser Gly Asp Leu Ser Pro Leu Ser Tyr Ile Ala Gly Ala			
W--> 2028	215	220	225	825
2029	atc acc ggt cac ccc gac gtc aag gtt cac gtt ttg cac gag gga acc			
2030	Ile Thr Gly His Pro Asp Val Lys Val His Val Leu His Glu Gly Thr			
W--> 2031	235	240	245	873
2032	gag aag atc atg ttt gcg cgc gag gcc atc tcg ctc ttt ggt ctc gag			
2033	Glu Lys Ile Met Phe Ala Arg Glu Ala Ile Ser Leu Phe Gly Leu Glu			
W--> 2034	250	255	260	930
2035	gca gtc g gtacgtcgag agtccctgact gcagtgagct gttcgagagt ctcccagttt			
2036	Ala Val			

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/939,408A

DATE: 01/23/2002
TIME: 17:03:26

Input Set : A:\500nscseq.txt
Output Set: N:\CRF3\01232002\I939408A.raw

W--> 2038	gctgactgcc ctttgttcat gcgattgcag tc ctc ggc ccg aag gag ggt ctc	983
2039	Val Leu Gly Pro Lys Glu Gly Leu	
W--> 2040	270	
2041	ggt ctg gtc aac gga acg gcc gtc tcc gcc tcg atg gcg acc ctc agt	1031
2042	Gly Leu Val Asn Gly Thr Ala Val Ser Ala Ser Met Ala Thr Leu Ser	
2043	275	280
2044	ctg cac gac tcg cac atg ctc tcg ctc ctc tcg cag gcc ttg acg gct	1079
2045	Leu His Asp Ser His Met Leu Ser Leu Leu Ser Gln Ala Leu Thr Ala	
2046	290	295
2047	ctc acg gtg gag gcc atg gtc ggc cag cag ggc tcg ttc gcg ccg ttc	1127
2048	Leu Thr Val Glu Ala Met Val Gly Gln Gln Gly Ser Phe Ala Pro Phe	
2049	305	310
2050	atc cac gac gtc tgc cgc ccg cac ccc ggc cag gtc gag gtc gcg cgc	1175
2051	Ile His Asp Val Cys Arg Pro His Pro Gly Gln Val Glu Val Ala Arg	
2052	325	330
2053	aac atc cgc acg ctc ctt tcc ggc tcg tcg ttt gcc gtt gag cac gag	1223
2054	Asn Ile Arg Thr Leu Leu Ser Gly Ser Ser Phe Ala Val Glu His Glu	
2055	340	345
2056	gag gag gtc aag gtc aag gac gac gag ggc att ctt cgc cag gac cgc	1271
2057	Glu Glu Val Lys Val Lys Asp Asp Glu Gly Ile Leu Arg Gln Asp Arg	
2058	355	360
2059	tac ccg ctc cgc acg tcg cct cag gttcgtcccc tctctctccc cttccctccg	1325
2060	Tyr Pro Leu Arg Thr Ser Pro Gln	
2061	370	375
2062	tccgaccggc gcgtcgagac ttacgttttg cgtatccag ttc ctc ggc ccg ctc	1379
2063	380	
2064	gtg gag gac atg atg cac gcc tac tcg act ctc tcg ctc gag aac aac	1427
2065	Val Glu Asp Met Met His Ala Tyr Ser Thr Leu Ser Leu Glu Asn Asn	
2066	385	390
2067	acg acg acc gac aac ccg ctc ctc gac gtc gag aac aag cag acc gcg	1475
2068	Thr Thr Thr Asp Asn Pro Leu Leu Asp Val Glu Asn Lys Gln Thr Ala	
2069	400	410
2070	cac ggc ggc aac ttc cag gcg tcg gct gtc tcg att tcg atg gag aag	1523
2071	His Gly Gly Asn Phe Gln Ala Ser Ala Val Ser Ile Ser Met Glu Lys	
2072	415	420
2073	acc agg tgcgtctctc gctgccttcg tactccgata ttgtgctgaa tgttcttctc	1579
2074	Thr Arg	
2075	430	
2076	ctgcagg ctc gca ctc gcc ctc atc ggc aag ctc aac ttc acg cag tgc	1628
2077	Leu Ala Leu Ala Leu Ile Gly Lys Leu Asn Phe Thr Gln Cys	
2078	435	440
2079	acc gag ttg ctc aac gct gcc atg aac cgc ggc ctg cct tcg tgc ctc	1676
2080	Thr Glu Leu Leu Asn Ala Ala Met Asn Arg Gly Leu Pro Ser Cys Leu	
2081	450	455
2082	gct gcc gag gac ccg tcg ctc aac tat cac ggc aag ggc ttg gac att	1724
2083	Ala Ala Glu Asp Pro Ser Leu Asn Tyr His Gly Lys Gly Leu Asp Ile	
2084	465	470
2085	cac atc gct gct tac gct tcg gag gtgagccgtc gacgttctcc gccgtcgctc	1778
2086		

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/939,408A

DATE: 01/23/2002
 TIME: 17:03:26

Input Set : A:\500nscseq.txt
 Output Set: N:\CRF3\01232002\I939408A.raw

2087	His Ile Ala Ala Tyr Ala Ser Glu	1833
2088	480	
2089	gtcccccttca ggcacccag gctgacttcc tttccctctg tag ctc ggc cac ctt	
2090	Leu Gly His Leu	
2092	gcc aac ccg gtc act acc ttc gtc cag ccc gca gag atg ggc aac cag	1881
2093	Ala Asn Pro Val Thr Thr Phe Val Gln Pro Ala Glu Met Gly Asn Gln	
2094	490	500
2095	gcc gtc aac tcg ctc gct ctc atc tcc gcg cgc cgc act gcc gag gcc	1929
2096	Ala Val Asn Ser Leu Ala Leu Ile Ser Ala Arg Arg Thr Ala Glu Ala	
2097	510	520
2098	aac gac gtc ctt tct ctc gtgcgttcgt gtgcgaatga gtcccgacgc	1977
2099	Asn Asp Val Leu Ser Leu	
2100	525	
2101	aatagcgact gactgcgcga tcttgagcag ctt ctc gcc tcg cac ctg tac tgc	2031
2102	Leu Leu Ala Ser His Leu Tyr Cys	
2103	530	535
2104	acg ctc cag gcc gtc gac ctc cgc gcg atg gag ctc gac ttc aag aag	2079
2105	Thr Leu Gln Ala Val Asp Leu Arg Ala Met Glu Leu Asp Phe Lys Lys	
2106	540	550
2107	cag ttc gac ccg ctt ctc ccg act ctc ctc cag cag cac ctc ggc act	2127
2108	Gln Phe Asp Pro Leu Leu Pro Thr Leu Leu Gln Gln His Leu Gly Thr	
2109	555	565
2110	ggc ctc gac gtc aac gca ctt gcg ctc gag gtc aag aag gcg ctc aac	2175
2111	Gly Leu Asp Val Asn Ala Leu Ala Leu Glu Val Lys Lys Ala Leu Asn	
2112	570	575
2113	aag cgt ctc gag cag acg acg acg tac gac ctc gag ccg cgc tgg cac	2223
2114	Lys Arg Leu Glu Gln Thr Thr Thr Tyr Asp Leu Glu Pro Arg Trp His	
2115	585	590
2116	gac gcc ttc tcg tac gcg acc ggc acc gtc gtc gag ctc ctc tcg tcc	2271
2117	Asp Ala Phe Ser Tyr Ala Thr Gly Thr Val Val Glu Leu Leu Ser Ser	
2118	600	605
2119	tcg ccc tct gcc aac gtc acc ctt act gcc gtc aac gcg tgg aag gtt	2319
2120	Ser Pro Ser Ala Asn Val Thr Leu Thr Ala Val Asn Ala Trp Lys Val	
2121	620	625
2122	gcc tcg gcc gag aag gcc atc tcg ctc acg cgc gag gtg cgc aac cgc	2367
2123	Ala Ser Ala Glu Lys Ala Ile Ser Leu Thr Arg Glu Val Arg Asn Arg	
2124	635	640
2125	ttc tgg cag acg ccg tct tcg cag gcg ccg gcg cac gca tac ctc tcg	2415
2126	Phe Trp Gln Thr Pro Ser Ser Gln Ala Pro Ala His Ala Tyr Leu Ser	
2127	650	655
2128	ccg cgc acg cgc gtc ctg tac tcg ttc gtg cgc gag gag ctc ggc gtg	2463
2129	Pro Arg Thr Arg Val Leu Tyr Ser Phe Val Arg Glu Glu Leu Gly Val	
2130	665	670
2131	cag gcg cgc cgc ggc gac gtg ttt gtc ggc gtg cag cag gag acg atc	2511
2132	Gln Ala Arg Arg Gly Asp Val Phe Val Gly Val Gln Gln Glu Thr Ile	
2133	680	685
2134	ggg agc aac gtc tcg cgc atc tac gag gcc atc aag gac ggc cgc atc	2559
2135	Gly Ser Asn Val Ser Arg Ile Tyr Glu Ala Ile Lys Asp Gly Arg Ile	
2136	700	705

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/939,408A

DATE: 01/23/2002

TIME: 17:03:26

Input Set : A:\500nscseq.txt

Output Set: N:\CRF3\01232002\I939408A.raw

2137	aac cac gtc ctc gtc aag atg ctc gcg taaggcccga gcaagcctcg	2606
2138	Asn His Val Leu Val Lys Met Leu Ala	
2139	715 720	
2140	cctagacgcc cgcctcacc caagaccagc ttttcgacgt cgtgtcgtgc caagaacgga	2666
2141	ctttcctcca tacacatgct gtcttactct ctgcgcgtca tcacgtctct cagttctttc	2726
2142	gtatcccgcg tctct	2741

1/23/02

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/939,408ADATE: 01/23/2002
TIME: 17:03:27Input Set : A:\500nscseq.txt
Output Set: N:\CRF3\01232002\I939408A.raw

L:61 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:79 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:197 M:283 W: Missing Blank Line separator, <220> field identifier
L:252 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:416 M:283 W: Missing Blank Line separator, <400> field identifier
L:444 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:1622 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1624 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1626 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1628 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1630 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1632 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1634 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1636 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1638 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1640 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1644 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1646 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1648 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1652 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1654 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1656 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1660 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1662 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1664 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1666 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1668 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1670 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1672 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1674 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1676 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1678 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1680 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1682 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1684 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1686 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1688 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1692 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1694 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1696 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1698 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1700 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1702 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:1704 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:1721 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:1724 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:1727 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:1730 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21

1/23/02

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/939,408ADATE: 01/23/2002
TIME: 17:03:27Input Set : A:\500nscseq.txt
Output Set: N:\CRF3\01232002\I939408A.raw

L:1733 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:1736 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:1738 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:1739 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:21
L:1741 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:1906 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:26
L:1907 M:283 W: Missing Blank Line separator, <220> field identifier
L:1909 M:283 W: Missing Blank Line separator, <400> field identifier
L:1914 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:27
L:1915 M:283 W: Missing Blank Line separator, <220> field identifier
L:1917 M:283 W: Missing Blank Line separator, <400> field identifier
L:1923 M:283 W: Missing Blank Line separator, <220> field identifier
L:1927 M:283 W: Missing Blank Line separator, <220> field identifier
L:1931 M:283 W: Missing Blank Line separator, <220> field identifier
L:1935 M:283 W: Missing Blank Line separator, <220> field identifier
L:1939 M:283 W: Missing Blank Line separator, <220> field identifier
L:1943 M:283 W: Missing Blank Line separator, <220> field identifier
L:1947 M:283 W: Missing Blank Line separator, <220> field identifier
L:1951 M:283 W: Missing Blank Line separator, <220> field identifier
L:1955 M:283 W: Missing Blank Line separator, <220> field identifier
L:1959 M:283 W: Missing Blank Line separator, <220> field identifier
L:1963 M:283 W: Missing Blank Line separator, <220> field identifier
L:1967 M:283 W: Missing Blank Line separator, <220> field identifier
L:1971 M:283 W: Missing Blank Line separator, <220> field identifier
L:1975 M:283 W: Missing Blank Line separator, <220> field identifier
L:1979 M:283 W: Missing Blank Line separator, <220> field identifier
L:1983 M:283 W: Missing Blank Line separator, <400> field identifier
L:1986 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:1986 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:6
L:1987 M:361 W: Invalid Split Codon, Sequence data for SEQ ID#: 28
L:1988 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:1991 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:1994 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:1997 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:2000 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:2003 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:2006 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:2010 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:2013 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:2016 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:2019 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:2022 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:2025 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:2028 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:2031 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:2034 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:2038 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:2040 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28
L:2148 M:283 W: Missing Blank Line separator, <220> field identifier

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/939,408A

DATE: 01/23/2002

TIME: 17:03:27

Input Set : A:\500nscseq.txt

Output Set: N:\CRF3\01232002\I939408A.raw

L:2151 M:283 W: Missing Blank Line separator, <220> field identifier
L:2155 M:283 W: Missing Blank Line separator, <400> field identifier
L:2296 M:283 W: Missing Blank Line separator, <220> field identifier
L:2299 M:283 W: Missing Blank Line separator, <220> field identifier
L:2303 M:283 W: Missing Blank Line separator, <220> field identifier
L:2307 M:283 W: Missing Blank Line separator, <220> field identifier
L:2311 M:283 W: Missing Blank Line separator, <220> field identifier
L:2315 M:283 W: Missing Blank Line separator, <220> field identifier
L:2319 M:283 W: Missing Blank Line separator, <220> field identifier
L:2323 M:283 W: Missing Blank Line separator, <220> field identifier
L:2327 M:283 W: Missing Blank Line separator, <220> field identifier
L:2331 M:283 W: Missing Blank Line separator, <220> field identifier
L:2335 M:283 W: Missing Blank Line separator, <220> field identifier
L:2339 M:283 W: Missing Blank Line separator, <220> field identifier
L:2343 M:283 W: Missing Blank Line separator, <220> field identifier
L:2347 M:283 W: Missing Blank Line separator, <220> field identifier
L:2351 M:283 W: Missing Blank Line separator, <220> field identifier
L:2355 M:283 W: Missing Blank Line separator, <220> field identifier
L:2359 M:283 W: Missing Blank Line separator, <220> field identifier
L:2363 M:283 W: Missing Blank Line separator, <220> field identifier
L:2367 M:283 W: Missing Blank Line separator, <220> field identifier
L:2371 M:283 W: Missing Blank Line separator, <220> field identifier
L:2375 M:283 W: Missing Blank Line separator, <220> field identifier
L:2379 M:283 W: Missing Blank Line separator, <220> field identifier
L:2383 M:283 W: Missing Blank Line separator, <220> field identifier
L:2387 M:283 W: Missing Blank Line separator, <220> field identifier
L:2391 M:283 W: Missing Blank Line separator, <220> field identifier